



07-20 09/889689 PCT

PATENT
DOCKET NO. BC45215
JC03 Rec'd PCT/PTC 19 JUL 2001

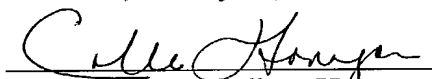
IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

BOX PCT/DO-EO

Commissioner for Patents
Washington, D.C. 20231

"Express Mail" mailing label number **ET 383 023 288 US**

I hereby certify that this paper is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" under 37 CFR § 1.10 on the date indicated above and addressed to: Commissioner of Patents, Washington, DC 20231, on July 19, 2001.


Colleen Hanagan

NATIONAL STAGE APPLICATION TRANSMITTAL LETTER
APPLICATION FILING UNDER 35 U.S.C. § 371

Transmitted herewith for filing is the patent application of:

Inventor(s)/Applicant(s):	Vinals-Bassols, Carlota
International Application No.:	PCT/EP00/00346
International Published Appln. No.:	WO 00/43509
International Filing Date:	17 January 2000
Priority Filing Dates:	19 January 1999, 29 January 1999, 01 February 1999 & 07 April 1999
Thirty Month Date:	19 July 2001
Title:	"NOVEL COMPOUNDS"

1. **THIS NEW APPLICATION IS A NATIONAL STAGE APPLICATION UNDER PCT, CHAPTER II WITH A REQUEST FOR EXAMINATION WITHOUT DELAY TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US).**

- ☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C. § 371;
- ☐ This is a **SECOND** or subsequent submission of items concerning a filing under 35 U.S.C. § 371.

2. ☒ This is an express request to begin national examination procedures (35 U.S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).

3. A proper Demand for International Preliminary Examination was made by the 19TH month from the earliest claimed priority date.

4. Enclosed items are required for filing under 37 CFR § 1.53(b) and § 1.494(b) or § 1.495(b):

- ☒ One copy of International Publication No. WO 00/43509
- (a) ☒ is transmitted herewith (**required only if not transmitted by the International Bureau**)
- (b) ☐ has been transmitted by the International Bureau
- (c) ☐ is not required, as the application was filed in the United States Receiving Office (RO/US)

Fees

☒ The basic national fee set forth in 37 CFR § 1.482 - International Preliminary Examination Fee not paid to USPTO but International Search Report prepared by the EPO or JPO - **\$860.00**

- ☐ Claims in Excess of 20 (@ \$18.00)
- ☒ Independent Claims in Excess of 3 (1 @ \$80.00)

5. Further enclosed are:

- ☒ One copy of International Preliminary Examination Report.
- ☒ One copy of International Search Report.
- ☐ One copy of Written Opinion.
- ☐ One copy of PCT Request as filed.
- ☐ One copy of Chapter II Demand as filed.

6. ☐ A translation of the International Application into English (35 U.S.C. § 371(c)(2))

7. ☐ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. § 371(c)(3))

8. ☐ A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. § 371(c)(3))

9. Still additional papers enclosed:

- ☐ Assignment with Assignment Recordation Form Cover Sheet
- ☐ Verification Statement Claiming Small Entity Status
- ☐ Declaration or oath is enclosed executed by the inventor
- ☐ An Information Disclosure Statement under 37 CFR § 1.97 and § 1.98
- ☒ Return Acknowledgment Postcard

09/889689

JC17 Rec'd PCT/PTO

19 JUL 2001 PATENT
DOCKET NO. BC45215

10. Nucleotide and/or Amino Acid Sequence Submission (if applicable, all necessary)

A. Enclosed are:

- (a) ☐ Computer Readable Copy of the Sequence Listing
(b) ☐ Paper Copy (identical to Computer Readable Copy) of the Sequence Listing

B. ☐ Enclosed is a paper copy of the Sequence Listing. This paper copy and a Computer Readable Form thereof are identical with the Computer Readable Form in another application of the Applicant which is fully identified as follows:

U.S. Application No.: @@

Filed: @@

Attorney Docket No.: @@

which is believed to comply with the rules set forth in 37 CFR § 1.821 et. seq. Applicants requests pursuant to 37 CFR § 1.821(e) that this Computer Readable Form be used in the present application. **Please TRANSFER the sequence listing from the parent to this application.**

C. ☐ Statement under 37 CFR § 1.821(f): **The information recorded in computer readable form is identical to the written Sequence Listing.**

D. ☐ Statement under 37 CFR § 1.821(g) (required when Sequence Listing not submitted at the time of filing under 35 U.S.C. §111(a)) or 37 CFR §1.821(f) (required when Sequence Listing not submitted at the time of filing under the Patent Cooperation Treaty): **The submission of the Sequence Listing includes no new matter.**

E. ☐ Amendment: Please enter the Sequence Listing into the application.

11. **Preliminary Amendment**

Prior to calculation of fees, kindly enter:

- ☒ Preliminary Amendment submitted herewith
☐ do not enter Preliminary Amendment

09/889689
JC17 Rec'd PCT/PTO 19 JUL 2001
PATENT
DOCKET NO. BC45215

12. The correspondence address for this application is the Customer No. provided below:

Insert Bar Code Label Here:



25308

PATENT TRADEMARK OFFICE

13. Fee payment being made at this time is enclosed:

* Basic filing fee (\$860.00)	860.00
* Claims in Excess of 20	00.00
* (@ \$18.00)	
Independent Claims in Excess of 3	80.00
(1 @ \$80.00)	
* Total Fees enclosed:	<u>\$940.00</u>

14. The Commissioner is hereby authorized to charge any fees which may be required, or credit any overpayment to Account No. 50-0258. This letter is filed in duplicate for accounting purposes.

Respectfully submitted,

E. A. Meade

Eric A. Meade
Registration No. 42,876
for
Allen Bloom
Registration No. 29,135
Attorney for Applicant

Date: July 19, 2001

DECHERT
Princeton Pike Corporate Center
PO Box 5218
Princeton, New Jersey 08543-5218
Allen Bloom (609) 620-3214
Eric A. Meade (609) 620-3248
Fax: (609) 620-3259
Attn.: Allen Bloom, Esq.
(609) 620-3214

International Application No.: PCT/EP00/00346
Attorney Docket No.: BC45215

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: Vinals-Bassols
Docket No.: BC45215
Serial No.: Unknown
Filed: Herewith
For: Novel Compounds

Group Art Unit No.: Unknown
Examiner: Unknown

PRELIMINARY AMENDMENT

Sir:

Please enter this Preliminary Amendment into the record of the above-identified patent application before examination and the calculation of fees.

In the Claims:

Please delete the claims of the application as filed in the PCT and substitute therefor:

6. An isolated polynucleotide comprising a nucleic acid sequence that is identical to one of SEQ ID NOs: 1-6, except that, over the entire length corresponding to one of SEQ ID NOs: 1-6, n_n nucleotides are substituted, inserted or deleted, wherein n_n satisfies the following expression

$$n_n \leq x_n - (x_n \cdot y)$$

wherein x_n is the total number of nucleotides in SEQ ID NO:1, y is 0.90, and wherein any non-integer product of x_n and y is rounded down to the nearest integer before subtracting the product from x_n .

7. The isolated polynucleotide of claim 6, wherein y is 0.95.

8. The isolated polynucleotide of claim 6, wherein the isolated polynucleotide consists of the nucleotide sequence of one of SEQ ID NOs: 1-6.

9. An isolated polynucleotide comprising a member selected from the group consisting of:

International Application No.: PCT/EP00/00346
Attorney Docket No.: BC45215

- (a) a polynucleotide comprising a nucleic acid sequence that is identical to one of SEQ ID NOs:1-6;
 - (b) a polynucleotide obtained by screening an appropriate library under stringent hybridization conditions, wherein stringent conditions comprise overnight incubation at 42° C. in a solution comprising: 50% formamide, 5×SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5× Denhardt's solution, 10% dextran sulfate, and 20 micrograms/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1× SSC at about 65° C; with a labeled probe having the sequence of one of SEQ ID Nos: 1-6, or a fragment thereof;
 - (c) a fragment of (a) or (b) which can be coupled to a suitable carrier, wherein the fragment is capable of raising an immune response which recognizes the polynucleotide of one of SEQ ID Nos: 1-6; or
- the full complement to the isolated polynucleotide.

10. A process for diagnosing a disease or a susceptibility to a disease in a subject related to a presence, expression, or activity of a polynucleotide of one of SEQ ID NOs:1-6 in a subject comprising analyzing for the presence or amount of the polynucleotide in a sample derived from the subject.

11. A process for diagnosing the presence of colon cancer or a susceptibility to colon cancer in a subject related to a presence, expression, or activity of a polynucleotide of one of SEQ ID NOs:1-6 in a subject comprising analyzing for the presence or amount of the polynucleotide in a sample derived from the subject.

International Application No.: PCT/EP00/00346
Attorney Docket No.: BC45215

REMARKS

Claims

Claims 1-5 have been canceled without prejudice or disclaimer of the subject matter therein. Applicant reserves the right to prosecute, in one or more patent applications, the canceled claims, the claims to non-elected inventions, the claims as originally filed, and any other claims supported by the specification.

New claims 6-11 have been introduced. No new matter is added.

Support

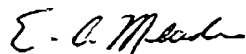
Support for the new claims is either obvious, or is as described in the text below. Particularly, support for the stringent hybridization conditions may be found, for example, at page 5, lines 28-32. Support for the recitation of sequence relatedness such as found in claim 6 may be found in the specification, for example, at page 26, line 15 through page 27, line 2. No new matter is added.

International Application No.: PCT/EP00/00346
Attorney Docket No.: BC45215

Closing Remarks

The Notice of Allowance is earnestly solicited.

Respectfully submitted,



Eric A. Meade
Registration No. 42,876
for
Allen Bloom
Registration No. 29,135
Attorney for Applicant

DECHERT
Princeton Pike Corporate Center
PO Box 5218
Princeton, New Jersey 08543-5218
Allen Bloom (609) 620-3214
Eric A. Meade (609) 620-3248
Fax: (609) 620-3259
Attn: Allen Bloom, Esq.
(609) 620-3214

Novel Compounds

FIELD OF THE INVENTION

The present invention relates to polynucleotides, (herein referred to as “CASB611
 5 polynucleotide(s)”, “CASB500 polynucleotide(s)”, “CASB501 polynucleotide(s)”,
 “CASB502 polynucleotide(s)”, “CASB505 polynucleotide(s)”, “CASB507
 polynucleotide(s)”), polypeptides encoded thereby (referred to herein as “CASB611”,
 “CASB500”, “CASB501”, “CASB502”, “CASB505”, and “CASB507” respectively or
 “CASB611 polypeptide(s)”, “CASB500 polypeptide (s)”, “CASB501 polypeptide (s)”,
 10 “CASB502 polypeptide (s)”, “CASB505 polypeptide (s)”, “CASB507 polypeptide
 (s)” respectively), recombinant materials and methods for their production. In another
 aspect, the invention relates to methods for using such polypeptides and polynucleotides,
 including the diagnostics and treatment of cancer and autoimmune diseases and other related
 conditions. In a further aspect, the invention relates to methods for identifying agonists
 15 and antagonists/inhibitors using the materials provided by the invention, and treating
 conditions associated with CASB611, CASB500, CASB501, CASB502, CASB505, or
 CASB507 polypeptide imbalance with the identified compounds. In a still further aspect,
 the invention relates to diagnostic assays for detecting diseases associated with inappropriate
 CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptide
 20 activity or levels.

BACKGROUND OF THE INVENTION

Polynucleotides and polypeptides of the present invention are believed to be important
 immunogens for specific prophylactic or therapeutic immunization against tumours, because
 25 they are specifically expressed or highly over-expressed in tumours compared to normal
 cells and can thus be targeted by antigen-specific immune mechanisms leading to the
 destruction of the tumour cell. They can also be used to diagnose the occurrence of tumour
 cells. Furthermore, their inappropriate expression in certain circumstances can cause an
 induction of autoimmune, inappropriate immune responses, which could be corrected
 30 through appropriate vaccination using the same polypeptides or polynucleotides. In this
 respect the most important biological activities to our purpose are the antigenic and
 immunogenic activities of the polypeptide of the present invention. A polypeptide of the
 present invention may also exhibit at least one other biological activity of a CASB611,

CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptide, which could qualify it as a target for therapeutic or prophylactic intervention different from that linked to the immune response.

5 Functional genomics relies heavily on high-throughput DNA sequencing technologies and the various tools of bioinformatics to identify gene sequences of potential interest from the many molecular biology databases now available. cDNA libraries enriched for genes of relevance to a particular tissue or physiological situation can be constructed using recently developed subtractive cloning strategies. Furthermore, cDNAs found in libraries of certain
10 tissues and not others can be identified using appropriate electronic screening methods.

High throughput genome- or gene-based biology allows new approaches to the identification and cloning of target genes for useful immune responses for the prevention and vaccine therapy of diseases such as cancer and autoimmunity.

15

DESCRIPTION OF THE INVENTION

Polynucleotides

In a first aspect, the present invention relates to CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polynucleotides. Such polynucleotides include
20 isolated polynucleotides comprising a nucleotide sequence which has at least 70% identity, preferably at least 80% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, to SEQ ID NOs:1 – 6 respectively over the entire length of SEQ ID NOs:1 - 6. In this regard, polynucleotides which have at least 97% identity are highly preferred, whilst those with at least 98-99% identity are more highly preferred, and
25 those with at least 99% identity are most highly preferred. Such polynucleotides include a polynucleotide comprising the polynucleotide of SEQ ID NOs:1 – 6 as well as the polynucleotide of SEQ ID NOs:1 - 6. Said polynucleotide can be inserted in a suitable plasmid or recombinant microorganism vector and used for immunization (see for example Wolff et. al., Science 247:1465-1468 (1990); Corr et. al., J. Exp. Med. 184:1555-1560
30 (1996); Doe et. al., Proc. Natl. Acad. Sci. 93:8578-8583 (1996)).

The invention also provides polynucleotides which are complementary to the above described polynucleotides.

WO 00/43509

PCT/EP00/00346

The invention also provides a polynucleotide comprising a nucleotide sequence that has at least 70% identity to a nucleotide sequence encoding a polypeptide of the invention over the entire coding region, encoded by a polynucleotide comprising the sequence contained in any one of SEQ ID Nos: 1 - 6; or a nucleotide sequence complimentary to said isolated
5 polynucleotide.

The invention also provides a fragment of a CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polynucleotide which when administered to a subject has the same immunogenic properties as the polynucleotide of SEQ ID NOs:1 - 6.

10

The invention also provides a polynucleotide encoding an immunological fragment of a CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptide as hereinbefore defined.

15 The nucleotide sequences of SEQ ID NO:1 (CASB611) and SEQ ID NO:2 show no homology with any known gene.. The nucleotide sequences of SEQ ID NO:3-6 show homology to intron portions of the chromosomal PBEF gene (Homo sapiens BAC clone RP11-22N19 from 7q22, complete sequence; ACCESSION : AC007032), SEQ ID NO: 3 and 5 (CASB501 and CASB505) being located between exon 8 and exon 7 of the PBEF
20 gene (nucleotides 67 484 to 67 966 and 68 107 to 68 583), and SEQ ID NO:4 (CASB502) and 6 (CASB507) being located downstream of exon 8 (from 69 274 to 69 732 and from 70 163 to 71 223).

25 CASB500 is located on chromosome 6 (accession number HSJ651N20).

25

Preferred polypeptides and polynucleotides of the present invention are expected to have, *inter alia*, similar biological functions/properties to their homologous polypeptides and polynucleotides. Furthermore, preferred polynucleotides of the present invention have at least one activity of SEQ ID NOs:1 - 6, as appropriate.

30

The present invention also relates to partial polynucleotide and polypeptide sequences which were first identified prior to the determination of the corresponding sequences of SEQ ID NOs:1 - 6.

Accordingly, in a further aspect, the present invention provides for an isolated polynucleotide which:

- 5 (a) comprises a nucleotide sequence which has at least 70% identity, preferably at least 80% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, even more preferably at least 97-99% identity to SEQ ID NO:7-12 over the entire length of SEQ ID NO:7-12;
- (b) has a nucleotide sequence which has at least 70% identity, preferably at least 80% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, 10 even more preferably at least 97-99% identity, to SEQ ID NO:1-6 over the entire length of SEQ ID NO:7-12; or
- (c) the polynucleotides of SEQ ID NO:7-12; or

15 The nucleotide sequences of SEQ ID NO:7-12 and the peptide sequence encoded thereby are derived from EST (Expressed Sequence Tag) sequences. It is recognised by those skilled in the art that there will inevitably be some nucleotide sequence reading errors in EST sequences (see Adams, M.D. *et al*, Nature 377 (supp) 3, 1995). Accordingly, the nucleotide sequence of SEQ ID NO:7-12 and the peptide sequence encoded therefrom are therefore subject to the same inherent limitations in sequence accuracy.

20 Polynucleotides of the present invention may be obtained, using standard cloning and screening techniques, from a cDNA library derived from mRNA in cells of human colon cancer, (for example Sambrook *et al*, Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring harbor Laboratory Press, Cold Spring harbor, N.Y. (1989)). Polynucleotides of 25 the invention can also be obtained from natural sources such as genomic DNA libraries or can be synthesized using well known and commercially available techniques.

When polynucleotides of the present invention are used for the recombinant production of polypeptides of the present invention, the polynucleotide may include the coding sequence 30 for the mature polypeptide, by itself; or the coding sequence for the mature polypeptide in reading frame with other coding sequences, such as those encoding a leader or secretory sequence, a pre-, or pro- or prepro- protein sequence, or other fusion peptide portions. For example, a marker sequence which facilitates purification of the fused polypeptide can be

encoded. In certain preferred embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz *et al.*, Proc Natl Acad Sci USA (1989) 86:821-824, or is an HA tag. The polynucleotide may also contain non-coding 5' and 3' sequences, such as transcribed, non-translated sequences, splicing and polyadenylation signals, ribosome binding sites and sequences that stabilize mRNA.

Polynucleotides which are identical or sufficiently identical to a nucleotide sequence contained in SEQ ID NOs:1 - 6, may be used as hybridization probes for cDNA and genomic DNA or as primers for a nucleic acid amplification (PCR) reaction, to isolate full-length cDNAs and genomic clones encoding polypeptides of the present invention and to isolate cDNA and genomic clones of other genes (including genes encoding paralogs from human sources and orthologs and paralogs from species other than human) that have a high sequence similarity to SEQ ID NO:1 - 6. Typically these nucleotide sequences are 70% identical, preferably 80% identical, more preferably 90% identical, most preferably 95% identical to that of the referent. The probes or primers will generally comprise at least 15 nucleotides, preferably, at least 30 nucleotides and may have at least 50 nucleotides. Particularly preferred probes will have between 30 and 50 nucleotides. Particularly preferred primers will have between 20 and 25 nucleotides. In particular, polypeptides or polynucleotides derived from sequences from homologous animal origin could be used as immunogens to obtain a cross-reactive immune response to the human gene.

A polynucleotide encoding a polypeptide of the present invention, including homologs from species other than human, may be obtained by a process which comprises the steps of screening an appropriate library under stringent hybridization conditions with a labeled probe having the sequence of SEQ ID NOs: 1 - 6 or a fragment thereof; and isolating full-length cDNA and genomic clones containing said polynucleotide sequence. Such hybridization techniques are well known to the skilled artisan. Preferred stringent hybridization conditions include overnight incubation at 42°C in a solution comprising: 50% formamide, 5xSSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10 % dextran sulfate, and 20 microgram/ml denatured, sheared salmon sperm DNA; followed by washing the filters in 0.1x SSC at about 65°C.

Thus the present invention also includes polynucleotides obtainable by screening an appropriate library under stringent hybridization conditions with a labeled probe having the sequence of SEQ ID NOS:1 – 6 or a fragment thereof.

- 5 The skilled artisan will appreciate that, in many cases, an isolated cDNA sequence will be incomplete, in that the region coding for the polypeptide is short at the 5' end of the cDNA.

There are several methods available and well known to those skilled in the art to obtain
10 full-length cDNAs, or extend short cDNAs, for example those based on the method of Rapid Amplification of cDNA ends (RACE) (see, for example, Frohman et al., PNAS USA 85, 8998-9002, 1988). Recent modifications of the technique, exemplified by the Marathon™ technology (Clontech Laboratories Inc.) for example, have significantly
15 simplified the search for longer cDNAs. In the Marathon™ technology, cDNAs have been prepared from mRNA extracted from a chosen tissue and an 'adaptor' sequence ligated onto each end. Nucleic acid amplification (PCR) is then carried out to amplify the 'missing' 5' end of the cDNA using a combination of gene specific and adaptor specific oligonucleotide primers. The PCR reaction is then repeated using 'nested' primers, that is,
20 primers designed to anneal within the amplified product (typically an adaptor specific primer that anneals further 3' in the adaptor sequence and a gene specific primer that anneals further 5' in the known gene sequence). The products of this reaction can then be analysed by DNA sequencing and a full-length cDNA constructed either by joining the product directly to the existing cDNA to give a complete sequence, or carrying out a
25 separate full-length PCR using the new sequence information for the design of the 5' primer.

Polypeptides

In a further aspect, the present invention relates to CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptides.

30

Further peptides of the present invention include isolated polypeptides encoded by a polynucleotide comprising the sequence contained in one of SEQ ID NOS:1 - 6.

The invention also provides an immunogenic fragment of a CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptide, that is a contiguous portion of the CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptide which has the same or similar immunogenic properties to the polypeptide encoded by a polynucleotide comprising the sequence contained in one of SEQ ID NOS:1 – 6. That is to say, the fragment (if necessary when coupled to a carrier) is capable of raising an immune response which recognises the CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptide. Such an immunogenic fragment may include, for example, the CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptide lacking an N-terminal leader sequence, a transmembrane domain or a C-terminal anchor domain.

The polypeptides or immunogenic fragment of the invention may be in the form of the “mature” protein or may be a part of a larger protein such as a precursor or a fusion protein. It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification such as multiple histidine residues, or an additional sequence for stability during recombinant production. Furthermore, addition of exogenous polypeptide or lipid tail or polynucleotide sequences to increase the immunogenic potential of the final molecule is also considered.

In one aspect, the invention relates to genetically engineered soluble fusion proteins comprising a polypeptide of the present invention, or a fragment thereof, and various portions of the constant regions of heavy or light chains of immunoglobulins of various subclasses (IgG, IgM, IgA, IgE). Preferred as an immunoglobulin is the constant part of the heavy chain of human IgG, particularly IgG1, where fusion takes place at the hinge region. In a particular embodiment, the Fc part can be removed simply by incorporation of a cleavage sequence which can be cleaved with blood clotting factor Xa. Furthermore, this invention relates to processes for the preparation of these fusion proteins by genetic engineering, and to the use thereof for drug screening, diagnosis and therapy. A further aspect of the invention also relates to polynucleotides encoding such fusion proteins. Examples of fusion protein technology can be found in International Patent Application Nos. WO94/29458 and WO94/22914.

The proteins may be chemically conjugated, or expressed as recombinant fusion proteins allowing increased levels to be produced in an expression system as compared to non-fused protein. The fusion partner may assist in providing T helper epitopes
 5 (immunological fusion partner), preferably T helper epitopes recognised by humans, or assist in expressing the protein (expression enhancer) at higher yields than the native recombinant protein. Preferably the fusion partner will be both an immunological fusion partner and expression enhancing partner.

10 Fusion partners include protein D from *Haemophilus influenza* B and the non-structural protein from influenzae virus, NS1 (hemagglutinin). Another immunological fusion partner is the protein known as LYTA. Preferably the C terminal portion of the molecule is used. Lyta is derived from *Streptococcus pneumoniae* which synthesize an N-acetyl-L-alanine amidase, amidase LYTA, (coded by the *lytA* gene {Gene, 43 (1986) page 265-
 15 272} an autolysin that specifically degrades certain bonds in the peptidoglycan backbone. The C-terminal domain of the LYTA protein is responsible for the affinity to the choline or to some choline analogues such as DEAE. This property has been exploited for the development of E.coli C-LYTA expressing plasmids useful for expression of fusion
 20 proteins. Purification of hybrid proteins containing the C-LYTA fragment at its amino terminus has been described {Biotechnology: 10, (1992) page 795-798}. It is possible to use the repeat portion of the Lyta molecule found in the C terminal end starting at residue 178, for example residues 188 - 305.

The present invention also includes variants of the aforementioned polypeptides, that is
 25 polypeptides that vary from the referents by conservative amino acid substitutions, whereby a residue is substituted by another with like characteristics. Typical such substitutions are among Ala, Val, Leu and Ile; among Ser and Thr; among the acidic residues Asp and Glu; among Asn and Gln; and among the basic residues Lys and Arg; or aromatic residues Phe and Tyr. Particularly preferred are variants in which several, 5-10, 1-5, 1-3, 1-2 or 1 amino
 30 acids are substituted, deleted, or added in any combination.

Polypeptides of the present invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced

polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

5

Vectors, Host cells, Expression Systems

Recombinant polypeptides of the present invention may be prepared by processes well known in the art from genetically engineered host cells comprising expression systems. Accordingly, in a further aspect, the present invention relates to an expression system which comprises a polynucleotide of the present invention, to host cells which are genetically engineered with such expression systems and to the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention.

15

For recombinant production, host cells can be genetically engineered to incorporate expression systems or portions thereof for polynucleotides of the present invention. Introduction of polynucleotides into host cells can be effected by methods described in many standard laboratory manuals, such as Davis et al., Basic Methods in Molecular Biology (1986) and Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989). Preferred such methods include, for instance, calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction or infection.

25

Preferably the proteins of the invention are coexpressed with thioredoxin in trans (TIT). Coexpression of thioredoxin in trans versus in cis is preferred to keep antigen free of thioredoxin without the need for protease. Thioredoxin coexpression eases the solubilisation of the proteins of the invention. Thioredoxin cocexpression has also a significant impact on protein purification yield, on purified-protein solubility and quality.

Representative examples of appropriate hosts include bacterial cells, such as *Streptococci*, *Staphylococci*, *E. coli*, *Streptomyces* and *Bacillus subtilis* cells; fungal cells, such as yeast

cells and *Aspergillus* cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, HEK 293 and Bowes melanoma cells; and plant cells.

5 A great variety of expression systems can be used, for instance, chromosomal, episomal and virus-derived systems, e.g., vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and
10 retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression systems may contain control regions that regulate as well as engender expression. Generally, any system or vector which is able to maintain, propagate or express a polynucleotide to produce a polypeptide in a host may be used. The appropriate
15 nucleotide sequence may be inserted into an expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook *et al.*, Molecular Cloning, A Laboratory Manual (supra). Appropriate secretion signals may be incorporated into the desired polypeptide to allow secretion of the translated protein into the lumen of the endoplasmic reticulum, the periplasmic space or the extracellular
20 environment. These signals may be endogenous to the polypeptide or they may be heterologous signals.

The expression system may also be a recombinant live microorganism, such as a virus or bacterium. The gene of interest can be inserted into the genome of a live recombinant
25 virus or bacterium. Inoculation and *in vivo* infection with this live vector will lead to *in vivo* expression of the antigen and induction of immune responses. Viruses and bacteria used for this purpose are for instance: poxviruses (e.g; vaccinia, fowlpox, canarypox), alphaviruses (Sindbis virus, Semliki Forest Virus, Venezuelan Equine Encephalitis Virus), adenoviruses, adeno-associated virus, picornaviruses (poliovirus, rhinovirus),
30 herpesviruses (varicella zoster virus, etc), *Listeria*, *Salmonella*, *Shigella*, BCG. These viruses and bacteria can be virulent, or attenuated in various ways in order to obtain live vaccines. Such live vaccines also form part of the invention.

Polypeptides of the present invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, ion metal affinity chromatography (IMAC) is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during intracellular synthesis, isolation and or purification.

10 **Vaccines**

Another aspect of the invention relates to a method for inducing , re-inforcing or modulating an immunological response in a mammal which comprises inoculating the mammal with a fragment or the entire polypeptide or polynucleotide of the invention, adequate to produce antibody and/or T cell immune response for prophylaxis or for therapeutic treatment of cancer and autoimmune disease and related conditions. Yet another aspect of the invention relates to a method of inducing, re-inforcing or modulating immunological response in a mammal which comprises, delivering a polypeptide of the present invention *via* a vector or cell directing expression of the polynucleotide and coding for the polypeptide *in vivo* in order to induce such an immunological response to produce immune responses for prophylaxis or treatment of said mammal from diseases.

A further aspect of the invention relates to an immunological/vaccine formulation (composition) which, when introduced into a mammalian host, induces, re-inforces or modulates an immunological response in that mammal to a polypeptide of the present invention wherein the composition comprises a polypeptide or polynucleotide of the invention or an immunological fragment thereof as herein before defined. The vaccine formulation may further comprise a suitable carrier. Since a polypeptide may be broken down in the stomach, it is preferably administered parenterally (for instance, subcutaneous, intramuscular, intravenous, or intradermal injection). Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation isotonic with the blood of the recipient; and aqueous and non-aqueous

sterile suspensions which may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use.

5

A further aspect of the invention relates to the *in vitro* induction of immune responses to a fragment or the entire polypeptide or polynucleotide of the present invention or a molecule comprising the polypeptide or polynucleotide of the present invention, using cells from the immune system of a mammal, and reinfusing these activated immune cells of the mammal for the treatment of disease. Activation of the cells from the immune system is achieved by *in vitro* incubation with the entire polypeptide or polynucleotide of the present invention or a molecule comprising the polypeptide or polynucleotide of the present invention in the presence or absence of various immunomodulator molecules.

10

A further aspect of the invention relates to the immunization of a mammal by administration of antigen presenting cells modified by *in vitro* loading with part or the entire polypeptide of the present invention or a molecule comprising the polypeptide of the present invention and administered *in vivo* in an immunogenic way. Alternatively, antigen presenting cells can be transfected *in vitro* with a vector containing a fragment or the entire polynucleotide of the present invention or a molecule comprising the polynucleotide of the present invention, such as to express the corresponding polypeptide, and administered *in vivo* in an immunogenic way.

15

20

The vaccine formulation of the invention may also include adjuvant systems for enhancing the immunogenicity of the formulation. Preferably the adjuvant system raises preferentially a TH1 type of response.

25

An immune response may be broadly distinguished into two extreme categories, being a humoral or cell mediated immune responses (traditionally characterised by antibody and cellular effector mechanisms of protection respectively). These categories of response have been termed TH1-type responses (cell-mediated response), and TH2-type immune responses (humoral response).

30

Extreme TH1-type immune responses may be characterised by the generation of antigen specific, haplotype restricted cytotoxic T lymphocytes, and natural killer cell responses. In mice TH1-type responses are often characterised by the generation of antibodies of the IgG2a subtype, whilst in the human these correspond to IgG1 type antibodies. TH2-type
5 immune responses are characterised by the generation of a broad range of immunoglobulin isotypes including in mice IgG1, IgA, and IgM.

It can be considered that the driving force behind the development of these two types of immune responses are cytokines. High levels of TH1-type cytokines tend to favour the
10 induction of cell mediated immune responses to the given antigen, whilst high levels of TH2-type cytokines tend to favour the induction of humoral immune responses to the antigen.

The distinction of TH1 and TH2-type immune responses is not absolute. In reality an
15 individual will support an immune response which is described as being predominantly TH1 or predominantly TH2. However, it is often convenient to consider the families of cytokines in terms of that described in murine CD4 +ve T cell clones by Mosmann and Coffman (*Mosmann, T.R. and Coffman, R.L. (1989) TH1 and TH2 cells: different patterns of lymphokine secretion lead to different functional properties. Annual Review of*
20 *Immunology, 7, p145-173*). Traditionally, TH1-type responses are associated with the production of the INF- γ and IL-2 cytokines by T-lymphocytes. Other cytokines often directly associated with the induction of TH1-type immune responses are not produced by T-cells, such as IL-12. In contrast, TH2- type responses are associated with the secretion of IL-4, IL-5, IL-6 and IL-13.

25 It is known that certain vaccine adjuvants are particularly suited to the stimulation of either TH1 or TH2 - type cytokine responses. Traditionally the best indicators of the TH1:TH2 balance of the immune response after a vaccination or infection includes direct measurement of the production of TH1 or TH2 cytokines by T lymphocytes *in vitro* after
30 restimulation with antigen, and/or the measurement of the IgG1:IgG2a ratio of antigen specific antibody responses.

Thus, a TH1-type adjuvant is one which preferentially stimulates isolated T-cell populations to produce high levels of TH1-type cytokines when re-stimulated with antigen *in vitro*, and promotes development of both CD8+ cytotoxic T lymphocytes and antigen specific immunoglobulin responses associated with TH1-type isotype.

5

Adjuvants which are capable of preferential stimulation of the TH1 cell response are described in International Patent Application No. WO 94/00153 and WO 95/17209.

3 De-O-acylated monophosphoryl lipid A (3D-MPL) is one such adjuvant. This is known from GB 2220211 (Ribi). Chemically it is a mixture of 3 De-O-acylated monophosphoryl lipid A with 4, 5 or 6 acylated chains and is manufactured by Ribi Immunochem, Montana. A preferred form of 3 De-O-acylated monophosphoryl lipid A is disclosed in European Patent 0 689 454 B1 (SmithKline Beecham Biologicals SA).

15 Preferably, the particles of 3D-MPL are small enough to be sterile filtered through a 0.22micron membrane (European Patent number 0 689 454). 3D-MPL will be present in the range of 10µg - 100µg preferably 25-50µg per dose wherein the antigen will typically be present in a range 2-50µg per dose.

20 Another preferred adjuvant comprises QS21, an Hplc purified non-toxic fraction derived from the bark of Quillaja Saponaria Molina. Optionally this may be admixed with 3 De-O-acylated monophosphoryl lipid A (3D-MPL), optionally together with an carrier.

The method of production of QS21 is disclosed in US patent No. 5,057,540.

25

Non-reactogenic adjuvant formulations containing QS21 have been described previously (WO 96/33739). Such formulations comprising QS21 and cholesterol have been shown to be successful TH1 stimulating adjuvants when formulated together with an antigen.

30 Further adjuvants which are preferential stimulators of TH1 cell response include immunomodulatory oligonucleotides, for example unmethylated CpG sequences as disclosed in WO 96/02555.

Combinations of different TH1 stimulating adjuvants, such as those mentioned hereinabove, are also contemplated as providing an adjuvant which is a preferential stimulator of TH1 cell response. For example, QS21 can be formulated together with 3D-MPL. The ratio of QS21 : 3D-MPL will typically be in the order of 1 : 10 to 10 : 1; preferably 1:5 to 5 : 1 and often substantially 1 : 1. The preferred range for optimal synergy is 2.5 : 1 to 1 : 1 3D-MPL: QS21.

Preferably a carrier is also present in the vaccine composition according to the invention. The carrier may be an oil in water emulsion, or an aluminium salt, such as aluminium phosphate or aluminium hydroxide.

A preferred oil-in-water emulsion comprises a metabolisable oil, such as squalene, alpha tocopherol and Tween 80. In a particularly preferred aspect the antigens in the vaccine composition according to the invention are combined with QS21 and 3D-MPL in such an emulsion. Additionally the oil in water emulsion may contain span 85 and/or lecithin and/or tricaprylin.

Typically for human administration QS21 and 3D-MPL will be present in a vaccine in the range of 1µg - 200µg, such as 10-100µg, preferably 10µg - 50µg per dose. Typically the oil in water will comprise from 2 to 10% squalene, from 2 to 10% alpha tocopherol and from 0.3 to 3% tween 80. Preferably the ratio of squalene: alpha tocopherol is equal to or less than 1 as this provides a more stable emulsion. Span 85 may also be present at a level of 1%. In some cases it may be advantageous that the vaccines of the present invention will further contain a stabiliser.

Non-toxic oil in water emulsions preferably contain a non-toxic oil, e.g. squalane or squalene, an emulsifier, e.g. Tween 80, in an aqueous carrier. The aqueous carrier may be, for example, phosphate buffered saline.

A particularly potent adjuvant formulation involving QS21, 3D-MPL and tocopherol in an oil in water emulsion is described in WO 95/17210.

The present invention also provides a polyvalent vaccine composition comprising a vaccine formulation of the invention in combination with other antigens, in particular antigens useful for treating cancers, autoimmune diseases and related conditions. Such a polyvalent vaccine composition may include a TH-1 inducing adjuvant as hereinbefore described.

5

Diagnostic Assays, disease monitoring, chromosomal localisation, and gene amplification

An important aspect of the invention, in addition to the polynucleotides themselves, also relates to the use of the polynucleotides of the present invention and oligonucleotides
10 derived from them, as diagnostic and monitoring reagents. Oligonucleotide fragments derived from the polynucleotides of the invention for use as probes or primers generally comprise at least 15 bases. Particularly preferred probes will have between 30 and 50 nucleotides. Particularly preferred primers will have between 20 and 25 nucleotides.

15 The identification of genetic or biochemical markers in blood, other biological fluids, faeces, or tissues that will enable the detection of very early changes along the carcinogenesis pathway will help in determining the best treatment for the patient and the efficacy of the treatment. Surrogate tumour markers, such as polynucleotide expression, can be used to diagnose different forms and states of cancer. The identification of expression levels of the
20 polynucleotides of the invention will be useful in both the staging of the cancerous disorder and grading the nature of the cancerous tissue. The staging process monitors the advancement of the cancer and is determined on the presence or absence of malignant tissue in the areas biopsied. The polynucleotides of the invention can help to perfect the staging process by identifying markers for the aggressivity of a cancer, for example the presence in
25 different areas of the body. The grading of the cancer describes how closely a tumour resembles normal tissue of its same type and is assessed by its cell morphology and other markers of differentiation. The polynucleotides of the invention can be useful in determining the tumour grade as they can help in the determination of the differentiation status of the cells of a tumour.

30

On the other hand, the polypeptide of the invention can be produced by stroma cells, in which cases, its specific expression or differential expression is a marker of disease conditions.

Differential expression

The diagnostic assays offer a process for diagnosing or determining a susceptibility to cancers, autoimmune disease and related conditions through diagnosis by methods comprising determining from a sample derived from a subject an abnormally decreased or increased level of polypeptide or mRNA. This method of diagnosis is known as differential expression. The expression of a particular gene is compared between a diseased tissue and a normal tissue. A difference between the polynucleotide-related gene, mRNA, or protein in the two tissues is compared, for example in molecular weight, amino acid or nucleotide sequence, or relative abundance, indicates a change in the gene, or a gene which regulates it, in the tissue of the human that was suspected of being diseased.

Decreased or increased expression can be measured at the RNA level. PolyA RNA is first isolated from the two tissues and the detection of mRNA encoded by a gene corresponding to a differentially expressed polynucleotide of the invention can be detected by, for example, in situ hybridization in tissue sections, reverse transcriptase-PCR, using Northern blots containing poly A+ mRNA, or any other direct or indirect RNA detection method. An increased or decreased expression of a given RNA in a diseased tissue or surrounding tissues compared to a normal tissue or in the absence of disease suggests that the transcript and/or the expressed protein has a role in the disease. Thus detection of a higher or lower level of mRNA corresponding to SEQ ID NOS 1-6 relative to normal level is indicative of the presence of cancer in the patient.

mRNA expression levels in a sample can be determined by generation of a library of expressed sequence tags (ESTs) from the sample. The relative representation of ESTs in the library can be used to assess the relative representation of the gene transcript in the starting sample. The EST analysis of the test can then be compared to the EST analysis of a reference sample to determine the relative expression levels of the polynucleotide of interest.

Other mRNA analyses can be carried out using serial analysis of gene expression (SAGE) methodology (Velculescu et. Al. Science (1995) 270:484) , differential display methodology (For example, US 5,776,683) or hybridization analysis which relies on the specificity of nucleotide interactions.

Alternatively, where the polynucleotide sequence encodes a polypeptide, the comparison could be made at the protein level. The protein sizes in the two tissues may be compared using antibodies to detect polypeptides in Western blots of protein extracts from the two
5 tissues. Expression levels and subcellular localization may also be detected immunologically using antibodies to the corresponding protein. Further assay techniques that can be used to determine levels of a protein, such as a polypeptide of the present invention, in a sample derived from a host are well-known to those of skill in the art. A raised or decreased level of polypeptide expression in the diseased tissue compared with the
10 same protein expression level in the normal tissue indicates that the expressed protein may be involved in the disease.

In the assays of the present invention, the diagnosis can be determined by detection of gene product expression levels encoded by at least one sequence set forth in SEQ ID NOS: 1 - 6.
15 A comparison of the mRNA or protein levels in a diseased versus normal tissue may also be used to follow the progression or remission of a disease.

Use of arrays for diagnosis

A large number of polynucleotide sequences in a sample can be assayed using
20 polynucleotide arrays. These can be used to examine differential expression of genes and to determine gene function. For example, arrays of the polynucleotide sequences SEQ ID NOS 1 - 6 can be used to determine if any of the polynucleotides are differentially expressed between a normal and cancer cell. In one embodiment of the invention, an array of oligonucleotides probes comprising SEQ ID NOS 1 - 6 nucleotide sequence or fragments
25 thereof can be constructed to conduct efficient screening of e.g., genetic mutations. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability (see for example: M.Chee et al., Science, Vol 274, pp 610-613 (1996)).

30 "Diagnosis" as used herein includes determination of a subject's susceptibility to a disease, determination as to whether a subject presently has the disease, and also the prognosis of a subject affected by the disease.

The present invention, further relates to a diagnostic kit for performing a diagnostic assay which comprises:

- (a) a polynucleotide of the present invention, preferably the nucleotide sequence of any one of SEQ ID NOs: 1 – 6, or a fragment thereof ;
- (b) a nucleotide sequence complementary to that of (a);
- (c) a polypeptide of the present invention, preferably the polypeptide encoded by a polynucleotide comprising the sequence contained in any one of SEQ ID Nos 1 – 6 or a fragment thereof; or
- (d) an antibody to a polypeptide of the present invention, preferably to the polypeptide encoded by a polynucleotide comprising the sequence contained in any one of SEQ ID Nos 1 – 6.
- (e) any specific ligand to a polypeptide of the present invention

The nucleotide sequences of the present invention are also valuable for chromosomal localisation. The sequence is specifically targeted to, and can hybridize with, a particular location on an individual human chromosome. The mapping of relevant sequences to chromosomes according to the present invention is an important first step in correlating those sequences with gene associated disease. Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found in, for example, V. McKusick, Mendelian Inheritance in Man (available on-line through Johns Hopkins University Welch Medical Library). The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes). The differences in the cDNA or genomic sequence between affected and unaffected individuals can also be determined. In addition, gene amplification at the genomic level could be assessed and correlated with the stage of disease.

Antibodies

The polypeptides of the invention or their fragments or analogs thereof, or cells expressing them, can also be used as immunogens to produce antibodies immunospecific for polypeptides of the present invention. The term “immunospecific” means that the antibodies

have substantially greater affinity for the polypeptides of the invention than their affinity for other related polypeptides in the prior art.

In a further aspect the invention provides an antibody immunospecific for a polypeptide
 5 according to the invention or an immunological fragment thereof as hereinbefore defined. Preferably the antibody is a monoclonal antibody.

Antibodies generated against polypeptides of the present invention may be obtained by administering the polypeptides or epitope-bearing fragments, analogs or cells to an animal,
 10 preferably a non-human animal, using routine protocols. For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler, G. and Milstein, C., Nature (1975) 256:495-497), the trioma technique, the human B-cell hybridoma technique (Kozbor *et al.*, Immunology Today (1983) 4:72) and the EBV-hybridoma
 15 technique (Cole *et al.*, Monoclonal Antibodies and Cancer Therapy, 77-96, Alan R. Liss, Inc., 1985).

Techniques for the production of single chain antibodies, such as those described in U.S. Patent No. 4,946,778, can also be adapted to produce single chain antibodies to polypeptides
 20 of this invention. Also, transgenic mice, or other organisms, including other mammals, may be used to express humanized antibodies.

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptide or to purify the polypeptides by affinity chromatography.
 25 The antibody of the invention may also be employed to prevent or treat cancer, particularly colon cancer, autoimmune disease and related conditions.

Another aspect of the invention relates to a method for inducing or modulating an immunological response in a mammal which comprises inoculating the mammal with a
 30 polypeptide of the present invention, adequate to produce antibody and/or T cell immune response to protect or ameliorate the symptoms or progression of the disease. Yet another aspect of the invention relates to a method of inducing or modulating immunological response in a mammal which comprises, delivering a polypeptide of the

present invention *via* a vector directing expression of the polynucleotide and coding for the polypeptide *in vivo* in order to induce such an immunological response to produce antibody to protect said animal from diseases.

- 5 It will be appreciated that the present invention therefore provides a method of diagnosing and/or treating abnormal conditions such as, for instance, cancer and autoimmune diseases, in particular, colon cancer, related to either a presence of, an excess of, or an under-expression of, any one of CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507.

10

Screening

The present invention further provides for a method of screening compounds to identify those which stimulate or which inhibit the function of any one of the CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptides. In general, agonists or
15 antagonists may be employed for therapeutic and prophylactic purposes for such diseases as hereinbefore mentioned. Compounds may be identified from a variety of sources, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. Such agonists, antagonists or inhibitors so-identified may be natural or modified substrates, ligands, receptors, enzymes, etc., as the case may be, of the polypeptide; or may be structural
20 or functional mimetics thereof (see Coligan *et al.*, Current Protocols in Immunology 1(2):Chapter 5 (1991)). Screening methods will be known to those skilled in the art. Further screening methods may be found in for example D. Bennett *et al.*, J Mol Recognition, 8:52-58 (1995); and K. Johanson *et al.*, J Biol Chem, 270(16):9459-9471 (1995) and references therein.

25

Thus the invention provides a method for screening to identify compounds which stimulate or which inhibit the function of the polypeptide of the invention which comprises a method selected from the group consisting of:

- (a) measuring the binding of a candidate compound to the polypeptide (or to the cells or
30 membranes bearing the polypeptide) or a fusion protein thereof by means of a label directly or indirectly associated with the candidate compound;

- (b) measuring the binding of a candidate compound to the polypeptide (or to the cells or membranes bearing the polypeptide) or a fusion protein thereof in the presense of a labeled competitor;
- (c) testing whether the candidate compound results in a signal generated by activation or inhibition of the polypeptide, using detection systems appropriate to the cells or cell membranes bearing the polypeptide;
- (d) mixing a candidate compound with a solution containing a polypeptide of claim 1, to form a mixture, measuring activity of the polypeptide in the mixture, and comparing the activity of the mixture to a standard; or
- (e) detecting the effect of a candidate compound on the production of mRNA encoding said polypeptide and said polypeptide in cells, using for instance, an ELISA assay.

The polypeptide of the invention may be used to identify membrane bound or soluble receptors, if any, through standard receptor binding techniques known in the art. Well known screening methods may also be used to identify agonists and antagonists of the polypeptide of the invention which compete with the binding of the polypeptide of the invention to its receptors, if any.

Thus, in another aspect, the present invention relates to a screening kit for identifying agonists, antagonists, ligands, receptors, substrates, enzymes, etc. for polypeptides of the present invention; or compounds which decrease or enhance the production of such polypeptides, which comprises:

- (a) a polypeptide of the present invention;
- (b) a recombinant cell expressing a polypeptide of the present invention;
- (c) a cell membrane expressing a polypeptide of the present invention; or
- (d) antibody to a polypeptide of the present invention;

which polypeptide is preferably that encoded by a polynucleotide comprising the sequence contained in any one of SEQ ID NOs:1 - 6.

30

It will be readily appreciated by the skilled artisan that a polypeptide of the present invention may also be used in a method for the structure-based design of an agonist, antagonist or inhibitor of the polypeptide, by:

- (a) determining in the first instance the three-dimensional structure of the polypeptide;
- (b) deducing the three-dimensional structure for the likely reactive or binding site(s) of an agonist, antagonist or inhibitor;
- (c) synthesising candidate compounds that are predicted to bind to or react with the deduced binding or reactive site; and
- (d) testing whether the candidate compounds are indeed agonists, antagonists or inhibitors.

Gene therapy may also be employed to effect the endogenous production of CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptides by the relevant cells in the subject. For an overview of gene therapy, see Chapter 20, Gene Therapy and other Molecular Genetic-based Therapeutic Approaches, (and references cited therein) in Human Molecular Genetics, T Strachan and A P Read, BIOS Scientific Publishers Ltd (1996).

Compositions and administration

Vaccine preparation is generally described in Pharmaceutical Biotechnology, Vol.61 Vaccine Design - the subunit and adjuvant approach, edited by Powell and Newman, Plenum Press, 1995. New Trends and Developments in Vaccines, edited by Voller et al., University Park Press, Baltimore, Maryland, U.S.A. 1978. Encapsulation within liposomes is described, for example, by Fullerton, U.S. Patent 4,235,877. Conjugation of proteins to macromolecules is disclosed, for example, by Likhite, U.S. Patent 4,372,945 and by Armor et al., U.S. Patent 4,474,757.

The amount of protein in each vaccine dose is selected as an amount which induces an immunoprotective response without significant, adverse side effects in typical vaccinees. Such amount will vary depending upon which specific immunogen is employed. Generally, it is expected that each dose will comprise 1-1000 μ g of protein, preferably 2-100 μ g, most preferably 4-40 μ g. An optimal amount for a particular vaccine can be ascertained by standard studies involving observation of antibody titres and other responses in subjects. Following an initial vaccination, subjects may receive a boost in about 4 weeks.

Definitions

"Isolated" means altered "by the hand of man" from the natural state. If an "isolated" composition or substance occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living animal is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein.

"Polynucleotide" generally refers to any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA including single and double stranded regions.

"Variant" refers to a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide, but retains essential properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques or by direct synthesis.

"Identity," as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between

- strings of such sequences. "Identity" and "similarity" can be readily calculated by known methods, including but not limited to those described in (Computational Molecular Biology, Lesk, A.M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D.W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; and Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., *SIAM J. Applied Math.*, 48: 1073 (1988). Preferred methods to determine identity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, the GCG program package (Devereux, J., et al., *Nucleic Acids Research* 12(1): 387 (1984)), BLASTP, BLASTN, and FASTA (Atschul, S.F. et al., *J. Molec. Biol.* 215: 403-410 (1990). The BLAST X program is publicly available from NCBI and other sources (BLAST Manual, Altschul, S., et al., NCBI NLM NIH Bethesda, MD 20894; Altschul, S., et al., *J. Mol. Biol.* 215: 403-410 (1990). The well known Smith Waterman algorithm may also be used to determine identity.
- The preferred algorithm used is FASTA. The preferred parameters for polypeptide or polynucleotide sequence comparison using this algorithm include the following:
- Gap Penalty:12
Gap extension penalty: 4
Word size: 2, max 6
- Preferred parameters for polypeptide sequence comparison with other methods include the following:
- 1) Algorithm: Needleman and Wunsch, *J. Mol Biol.* 48: 443-453 (1970)
Comparison matrix: BLOSSUM62 from Hentikoff and Hentikoff, *Proc. Natl. Acad. Sci. USA.* 89:10915-10919 (1992)
Gap Penalty: 12
Gap Length Penalty: 4

EXAMPLES

Example 1

Subtractive cDNA cloning of colon tumour-associated antigen (TAA) candidates.

5 Subtractive cDNA libraries are produced using standard technologies. Briefly, total RNA is extracted from frozen (-70°C) colon tumour and matched normal colon samples using the TriPure reagent and protocol (Boehringer). Target RNA is prepared by pooling total RNA from three tumour samples (30 µg each). Driver RNA is prepared by pooling total RNA from three matched normal colon samples (10 µg each) and total RNA from
10 seven normal tissues other than colon (brain, heart, kidney, liver, bladder, skin, spleen; 10 µg each). Total RNA from non-colon normal tissues is purchased from InVitrogen.

Messenger RNA is purified from total RNA using oligo-dT magnetic bead technology (Dyna) and quantified by spectrofluorimetry (BioRad).

15

Target and driver mRNA are reverse transcribed into cDNA using one of two strategies: 1) Target sequences for PCR oligonucleotides are introduced onto the ends of the newly synthesised cDNA during reverse transcription using the template switching capability of reverse transcriptase (ClonTech SMART PCR cDNA synthesis kit). 2) Target
20 and driver mRNA are reverse transcribed into cDNA using an oligo-dT primer and converted to double-strand cDNA; the cDNA is cleaved with RsaI and linkers for PCR amplification are ligated onto the extremities of the cDNA fragments.

In both cases, target and driver cDNA are amplified by long range PCR (ClonTech
25 SMART PCR Synthesis Kit and Advantage PCR Polymerase Mix) and used as starting material for subtractive cloning. For amplification, cycling conditions and optimisation of the number of PCR cycles are as described in the Advantage PCR protocol.

Two subtractive cloning strategies are used: ClonTech PCR SELECT (see
30 ClonTech kit protocol and N. Gurskaya et al. 1996. Analytical Biochemistry: 240, 90) and cRDA (M. Hubank and D. Schatz. 1994. Nucleic Acids Research: 22, 5640). When the PCR SELECT protocol is used, the primary PCR SELECT subtraction products are submitted to

a supplementary round of cRDA subtraction. When the cRDA protocol is used, two consecutive cycles of cRDA subtraction are performed.

In each case the products of both cycles of subtraction are cloned into pCR-TOPO (InVitrogen) and transformed into E. coli to produce a subtracted cDNA plasmid library.

An alternative strategy is also followed: subtraction of normal colon sequences and sequences from non-colon normal tissues are subtracted in separate hybridisations. In this case, target and driver RNA are assembled for the first subtraction as above with the exception that non-colon RNA is left out of the driver pool and amounts of normal colon are increased to 10 µg. Preparation of target and driver cDNA and subtractive hybridisation are performed as described above. A second subtraction is then performed on the products of the first subtraction, but the driver is now composed of a pool of normal colon and normal non-colon mRNA from the seven normal tissues.

15

Differential Screening of cDNA arrays.

Identification of tumour-associated genes in the subtracted cDNA library is accomplished by differential screening.

Total bacterial DNA is extracted from 100 µl over-night cultures. Bacteria are lysed with guanidium isothiocyanate and the bacterial DNA is affinity purified using magnetic glass (Boehringer). Plasmid inserts are recovered from the bacterial DNA by Advantage PCR amplification (Clontech). The PCR products are dotted onto two nylon membranes to produce high density cDNA arrays using the Biomek 96 HDRT tool (Beckman). The spotted cDNA is covalently linked to the membrane by UV irradiation. The first membrane is hybridised with a mixed cDNA probe prepared from the tumour of a single patient. The second membrane is hybridised with an equivalent amount of mixed cDNA probe prepared from normal colon of the same patient. The probe cDNA is prepared by PCR amplification as described above and is labelled using the AlkPhos Direct System (Amersham).

Hybridisation conditions and stringency washes are as described in the AlkPhos Direct kit. Hybridised probe is detected by chemiluminescence. Hybridisation intensities for each cDNA fragment on both blots (see figure 1) are measured by film densitometry or direct measurement (BioRad Fluor-S Max). The ratio of the tumour to normal hybridisation

20
25
30

WO 00/43509

PCT/EP00/00346

intensities (T/N) is calculated for each gene to evaluate the degree of over-expression in the tumour. Genes which are significantly over-expressed in colon tumours are followed-up. Significance is arbitrarily defined as one standard deviation of the T/N frequency distribution. Differential screening experiments are repeated using RNA from multiple patient donors (>18) to estimate the frequency of over-expressing tumours in the patient population.

In addition, the DNA arrays are hybridised with mixed cDNA probes from normal tissues other than colon (see list above) to determine the level of expression of the candidate gene in these tissues.

Example 2

Real-time RT-PCR analysis

Real-time RT-PCR (U. Gibson. 1996. Genome Research: 6,996) is used to compare mRNA transcript abundance of the candidate antigen in tumour and normal colon tissues from multiple patients. In addition, mRNA levels of the candidate gene are re-evaluated by this approach in a panel of normal tissues.

Total RNA is extracted from snap frozen colon tissue biopsies using TriPure reagent (Boehringer). Total RNA from normal tissues is from InVitrogen as above. Poly-A⁺ mRNA is purified from total RNA after DNAase treatment using oligo-dT magnetic beads (Dynal). Quantification of the mRNA is performed by spectrofluorimetry (BioRad) using SybrII dye (Molecular Probes). Primers for amplification are designed with the Perkin-Elmer Primer Express software using default options for TaqMan amplification conditions.

Real-time reactions are assembled according to standard PCR protocols using 2 ng of purified mRNA for each reaction. SybrI dye (Molecular Probes) is added at a final dilution of 1/75000 for real-time detection. Amplification (40 cycles) and real-time detection is performed in a PE 7700 system. Ct values are calculated using the 7700 Sequence Detector software for the tumour (CtT) and normal (CtN) samples of each patient. The difference between Ct values (CtN-CtT) is a direct measure of the difference in transcript levels between the tumour and normal tissues. As Ct values are log-linearly

related to copy number and that the efficiency of PCR amplification under the prevailing experimental conditions is close to the theoretical amplification efficiency, $2^{(Ct_N - Ct_T)}$ is an estimate of the relative transcript levels in the two tissues (i.e. fold mRNA over-expression in tumor). The percentage of over-expressing patients and the average level of mRNA over-expression in the tumours of these patients is calculated from the data set of patients.

The values indicated in table 1 are "Colon range Ct" and "colon mean Ct", representing the range and mean of Ct values for colon sample pairs, to assess the global mRNA abundance, 40 corresponding to the maximum Ct and thus to an absence of amplification.

TABLE 1: Expression of CASB genes in colon cancer and normal colon

Gene name	relevant /tested patients	colon range Ct	colon mean Ct
CASB611	15	20-40	28.4
CASB500	18	17-26.5	22
CASB501	17	20.5-40	25.7
CASB502	18	31.4-37.3	25.6
CASB505	17	22.6-40	29
CASB507	9	29.5-40	33.8

TABLE 2:

Gene name	Patients over-expressing CASB in colon tumours (%)	Average level of over-expression in colon tumours (fold)
CASB500	94%	5
CASB501	71%	5
CASB502	61%	7
CASB505	76%	18
CASB507	78%	7

In addition, Ct values for 12 normal tissues are reported. The tissues tested are bladder, brain, breast, cervix, heart, kidney, liver, lung, oesophagus, placenta and uterus. The range and mean of the Ct value is an indicator of the specificity of tissue expression. Thus, 2 types of sequences are described, classified into categories A and B. A sequence

of category A has a colon-specific expression. A sequence of category B is overexpressed in colon cancer.

5 TABLE 3: Expression of CASB genes in normal tissues

Gene name	Normal Tissues	NT Range Ct	NT mean Ct	class	Comments
CASB611	12	22-40	33	A	High only in rectum
CASB500	12	19-26.5	22.5	B	
CASB501	12	22.5-28.5	25.3	B	
CASB502	12	21.5-28.5	23	B	
CASB505	12	14-35	26.5	B	High in lung
CASB507	12	33.5-38.5	35	B	

NT range Ct and mean Ct: Ct values (range and mean) in normal tissues

10

Example 3

Identification of the full length cDNA sequence

- Colon tumour cDNA libraries are constructed using the Lambda Zap system (Stratagene) from 2 µg of polyA+ mRNA as described in the supplied protocol. 1.5 x10⁶ independent phage are plated for each screening of the library. Phage plaques are transferred onto nylon filters, hybridised using a cDNA probe labelled with AlkPhos Direct (Amersham Pharmacia) and positive phage are detected by chemiluminescence. The positive phage are excised from the agar plat, eluted in 500µl SM buffer and confirmed by gene-specific PCR.
- 20 Eluted phage are converted to single strand M13 bacteriophage by *in vivo* excision. The bacteriophage is then converted to double strand plasmid DNA by infection of *E. coli*. Infected bacteria are plated and submitted to a second round of screening with the cDNA probe. Plasmid DNA is purified from positive bacterial clones and submitted to Southern blot analysis to estimated the size of the cDNA inserts. CDNA inserts from multiple
- 25 independent clones are sequenced on both strands.

- When the full length gene cannot be obtained directly from the cDNA library, missing sequence is isolated using RACE technology (Marathon Kit, ClonTech.). This approach relies on reverse transcribing mRNA into double strand cDNA, ligating linkers onto the ends
- 30 of the cDNA and amplifying the desired extremity of the cDNA using a gene-specific

SEQUENCE INFORMATION

SEQ ID 1:

CTGCATTGAGTTTATTTATTTATTTATTACTTGTTTGGTGGATGCTCTTTCCAGCAGACTCATGTCAGGAG
5 GCTAGGAAGGGATTCCCTTGGGGACCACTGGATGCTGGTAGTTAAATGCCAGGAGGCTGAATGGACCTGAAGA
TGGAGGAGACTCTGCAGTCTTGGTCAGCCACCCTTGGGTGCTTGCCACCCTGCACTCTAGCAGGATTGATGG
TCTCTGGATTTGTAGCTGTGACCGGTGATGGTGGAATGCTCGGTGGTTTGCCTGGAGAGGCCACATGGTG
GCGACTGAGGCCCTGTGGGGTGAGGTTGGCTCATGATAGCTCTGAAAGTTGATGGCACAATTGAGACAAGGA
CGTGGAGTTCTGGAACTTTCCCAGGGTCCTTGAGCCCCGAGATGAAGCCCTCCAACATTGCAGCTCACTGCT
10 GTGAAGGCTGAAAGGGGAGACAGGAGGCTGCTGGGTACCTCAGGTGCCAGGCATGAGGTTCTCCGCTGTCTG
GAGTCTTGCTTTGGGAGTGCTACACCTGTGGGACTCCCCAGGCACCAACCCAAAGGCGGGGCTCTGGCTCAAG
CTTGATGGTTTACAGAGCCTCACATGATGTCCTTTTCAAGTTTGTGAGTTGGGGCTAGGGACTCGGTGGCA
GCAGAGCCTGGTTCTCCTGCCTGACCAGGCCGTTGGCAGGACGCCTCTTCCCCTCTGTGGAGAAGATGGAGA
AGCAGAGAAGCAGGATACAGAGCAGGCTGGAAGGACTAGAAGCCTCATGGTGAGAAGGCAGATTTTCCCTC
15 GTTCCGAAT

SEQ ID NO 2:

TTTTTTAGGGTTCAGTTCAGCTGATTTTATTTCTTCTCAAAAAAAGTTATTTACAGAAGGTATATATCAA
CAATCTGACAGGCAGTGAAGTTCAGATGATTAGCTGGCATGATTTTTCTTTTTTCCCCCAAACATTGTT
20 TTTGTGGCCTTGAATTTTAAGACAAATATTCTACACGGCATATTGCACAGGATGGATGGCAAAAAAAGTTT
AAAAACAAAAACCCCTTAACGGAAGTGCCTTAAAAAGGCAGACGTCTAGTGCCTGTATGTTATATTAACA
TACATACACACAATCTTTTTGCTTATTATAATACAGACTTAAATGTACAAAGATGTTTTCCACTTTTTTCAA
TTTTTAAACACAACAGCTATAAACCTGAACACATATGCTATCATCATGCCATAAGACTAAAACAATTATATT
TAGCGACAAGTAGAAAGGATTAAATAGTCAATACAAGAAATGAAAAACGCAGTACATAGTGTGCGAACTCA
25 AATCGGCATTTAGATAGATCCAGTGGTTTAAACGGCACGTTTTTGCTTATAAAAAAAGTGCAAAAAAGATGT
GGTTTACAAGTTAAAGCTACAGAATCCCTTTTTGCTGTAATTGCACCAAGTTTTAAAGCCTCTGGACAGAGCA
GTATTTGCTTTAAAGCTTTGTTTTCTTAAAGCTTACAGTGTGGGCTAATCTCCTCCCCTTTTTACAAG
ACGGGGGCCGGAGGGTGGACACTGGTGGCAGGTTAAGGGATACTGTCACTTTAAAGAAGCCTGCAGATTGAAG
TGTAACATGGAGAAATTAGGGGCTGATTTTTTAAAGTGTGTGAGATATTAACCAGCCGCCCTGTTATAAAA
30 TCAGGAAATCCAAACAGCGATTTACACCGATTACACCCCTTTATATATTTTTTACAAAAATACACTGAGA
AAATAATCAAACGTTTTTCATCTCTCTTGTCTTTTTTTGTTTTTTTAAAGTGTCAAAAGTCTACATTTAAATA
TAAAAAATTAAAGTTAAAGTCTAGCCCTTCAGTGAAGGAGACGTAAATGGCGTGGGTAACAACAACACTAC
CAAAAAAAGAAAAAAGAAAAAAGAAAAAAGGAAGGAATAAAGAAATAAAGGAAGTAAAAAGAAAG
GAAAGAAAAAAGGGACAAAAAAGAAAAAATATGTTTGGCCAGTATAAATACGTCCACATATAAATGGCAT
35 CTGATTACATTTACAAGGAAAAAGAAATACGAGGATGGAGCATCGGTGAAGGAAAAAACACGTCTTCTCAT
TTACACCTATAAGGAATAAACACACACACTGAGAAAAATTTGGCCCTGAATTGTTTTTTTAAAGTCCAGCA
CAGATTTGAGTTGCGTTTGAATCCTTTAAAGAGTTAAGAATGAAAAAAGCTGGTGATAATTTGTTGTAGG
AATCAAACATAGCGCCATCTATCTGCTTTTTATATTATCCTACACTATTTTAAAAACTGCTCAACAGTCTTA

WO 00/43509

PCT/EP00/00346

TACAGAAATCTTTAAAAGATAGACAGGATAACATGCTATATTAACCCCACCATTGAAATAATCCAACACCAT
CACGATTCCGATTAAGAGAAGNAAAAAAT

SEQ ID NO: 3

5 GAGAAATTTTAATTTGGTGTACAAAACCTTTTAGTAGTTAGTGAATGAGTAAGATGATTTGTGTATTGTGCAG
TTTGGCTTTAAGAGCAAAGGAGAACTTAGAGACTATGGGTTCTAGATGCTGTGTCTGCCTTTCACAGTTTCT
CTGAACCTTTGTGATTGAAC TGGATT CATATTGGAGTTATATATCAGACAAAAACCAAAGTCATTGGGCACAA
GCAGAATAAGTCCAAATTTAAAAGTATAATCATGAGTTTGTAAATAATGCTTTATATACTTCTTTAAAATAT
10 ACTTATTTAATATTAAGGATTTTTTTTTTCCCTTGAATCCATATAATTTGAAAGGTCTAAGGGGAAAAACTTGT
TTTGAGCCACTTTATGATGTGAATTCACCTTCATAGGGAAATAACATGCTTGGTTTTAATGTACTTTAACCC
TTTTAAAGTACACTGGACCAAATTTTAGAGAGCATATACATAGTGTTATAAAAT

SEQ ID NO: 4

15 CAGTTCCTTGAAATCTGAAATCCNGTTTCTGAAATCTTGAATTGCCTCCCANTTGGCCATTTTCCCTTTGAGC
CAAAAGTTTCAGATTTTGGAGCATTTTGGATTGGGATGCTCAACCTGTGTAAATATTTGTTAATATATAA
TGTAATAATTTTAAAAGATCTCAAACAATTAATTTTGAATAATACCATTAAATTATGCCAGCATTGGTCATT
CTCCCTAATTTTCTAGTCCGCCTCTAATCTGTTCTCCTCAGTAATAATAAAAAGATAGCACATTCATTTATT
20 ATAAAGCCCCAAGTAGGTATAAGTTTAACTCATAACTGTTTGAAGAGTCAGCTTTATGTCTGAAATAATCAC
TTTATACAATGT
TATACCTGGGGTAGTGCAGCATGCACATTTCAAGATTGTAAGACTACTACCCGAATCCAAATTGAGAGAG
AACCTGTTTATATAAAAACCTATTAGTGAACACTCTTCTTAACAGCTACAGAGGCTGCCAGTAGCTATGATGT
25 TAATTCCTCTACTTCATGCATCTGTAGGCTAAATGGAAAATTAAGATATGACTGAAGAGAAGCTCCTCTTCAT
ATTCAGGTCTGCAACTTTTCTCTTTAAAAGGAGAGAATTAAGATCAATGAATACCAATACTGGAGAACTTT
ACAACGTATGAATTAACCTCTACAGTCAAGGGGTCACTACTTATAACGTTGTATCAGTAAGTCACCTAACGA
TCTACTTGTATCATGTCTAGCAAACCTTCACTGTCCCCATCTTACTTGCTTTCTCCATGACACCTGAATTTTC
CTGATGCTTTCTTTCTCTTGACTTCAATGACATCTCCCATGTTTACCATTTTTCCCTATTCTATGGTTC
30 CCTTTTCAGTCTTCCAGATGCTGTGCTCTCCTTCTTTAATCTAACCCTGTCAGATGGCAATATTTCTCATAG
TTCTGACTTAGGCACCTTGTCTTCTGTAACATACCTTGTGTGAAACACTCAAAAGCTT

SEQ ID NO: 5

ACATTTTTTTATTACGTAACAGTAAAAACATGAAACCTATCAAGTTTGCTAAAACCAAATATTGCAGGCTAAG
ACAAGACTTTTCAGATTGGCCTATCATTTAATAGAATAAGAAAAGACCTCCTTAATTCGTTTCCCTTTCTTA
35 AGGGGTAAAGTATCATCTGCTCATTAACCTGACAAATGGTGTGGTCAAGTAGCAAGGCTGCAGCCAAAAGACCT
ACACTGGATATCTGACTCCTTATCTTTTCATGTTGTGTGACCTTGAGCAAGTTACTTTAACCTTTTAAAGCCT
CTGNTTTTTTCTTCTATAACAGAGACAGTAATATTATCACTGGTTTTTTTTTGNAGAGATTAATGGGCGTAT
CTTATGTAAATCACCTAAAATATTTTTCTAGGTAGTAGGCAAAATAAGCCTGCATTACTACTTTTGGGA
40 GAAGATAACCAGGTCCTTGGACATTACCACCAAAGGTACCCTG

SEQ ID NO: 6

CCGACGTCGACTATCCAAGATGTACTATTATTGGCATATAGTTTCAAAAATTCACAGAAGGGAGCCAGGTGC
TCTCATTGCCTAACAAAATAATGGAAATATGTATTTCATTCTAACATCTTGACATACAGTATAAAGGGCACTC
45 AGCAAGTGCTTTTTTAGTTAGACTGATTTTAAATGAGTAGATTCAGGAATACCAGAAAACGCACACACACACA
GATAATTGGTTGAAAATACGAGGTTAGAACTTAACCATGAAGAGGAGTAAGGCTACTTAAAGCGTTAAAA
ACTAATTGGGTAAAGGTATGGTTGACCCAGCTACTTCATTTGCCTTACGATGTATATTTCATTAACTAAGTCA
CTCACCTTCTCTGTGATGTTGATGGTTTGGTATAGTAAACATGAGATATGATTAAAGGTGATTACAGGGATAG
50 ATCAAGTGTCTGCCTAAGTAAATCTGGGNTTTCATTTTTTTTTCTAGACATCTTGGA

SEQ ID NO: 7

TTGAGTTTATTTATTTATTTATTACTTGTTTGTCTGGATGCTCTTTCCAGCAGACTCATGTCAGGAGGCTA
GGAAGGGATTCTTTGGGGACCACTGGATGCTGGTAGTTAAATGCCAGGAGGCTGAATGGACCTGAAGATGGA
55 GGAGACTCTGCAGTCTTGGTCAGCCACCCTTGGGTGCTTGCCACCCTGCACTCTAGCAGGATTGATGGTCTC
TGGATTGTAGCTGTGACCGGTCATGGTGGAAATGCTCGGTGGTTTGCAGTGGAGAGGCCACATGGTGGCGA

WO 00/43509

PCT/EP00/00346

CTGAGGCCCTGTGGGGTGAGGTTGGCTCATGATAGCTCTGAAAGTTGATGGCACAATTGAGACAAGGACGTG
GAGTTCTGGAACCTTTCCAGGGTCCTTGACAGCCGAGATGAAGCCCTCCAACATTGCAGCTCACTGCTGTGA
AGGCTGAAAGGGGAGACAGGAGGCTGCTGGGTACCTCAGGTGCCAGGCATGAGGTTCTCCGCTGTCTGGAGT
CTTGCTTTGGGAGTGCTACACCTGTGGGACTCCCCAGGCACCACCCAAAGGCGGGGCTCTGGCTCAAGCTTG
5 CATGGTTTACAGAGCCTCACATGATGTCCTTTTCAGGTTTGTGAGTTGGGGCTAGGGACTCGGTGGCAGCAG
AGCCTGGTTCTCCTGCCTGACCAGGCCGTTGGCAGGACGCCTCTTCCCCTCTGTGGAGAAGATGGAGAAGCA
GAGAAGCAGGATACAGAGCAGGCTGGAAAGGACTAGAAGCCTCATGGTGAGAAGGCAGATTTTCCCTCGTTC
CGAAT

10 SEQ ID NO: 8

TTTTTTAGGGTTCAGTTCAGCTGATTTTATTTCTTCTCAAAAAAGTTATTTACAGAA
GGTATATATCAACAATCTGACAGGCAGTGAACCTGACATGATTAGCTGGCATGATTTTTT
CTTTTTTTTCCCCCAAACATTGTTTTTGTGGCCTTGAATTTTAAGACAAATATTCTACAC
15 GGCATATTGCACAGGATGGATGGCAAAAAAAGTTTAAAAACAAAACCCTTAACGGAAC
TGCCTTAAAAAGGCAGACGTCCTAGTGCCTGTCATGTTATATTAAACATACATACACACA
ATCTTTTTTGCTTATTATAATACAGACTTAAATGTACAAAGATGTTTTCCACTTTTTTCAA
TTTTTAAACACAACAGCTATAAACCTGAACACATATGCTATCATCATGCCATAAGACTAA
ACAATTATATTTAGCGACAAGTAGAAAAGGATTAAATAGTCAAATACAAGAATGAAAAAC
20 GCAGTACATAGTGTGCGGAACTCAAATCGGCATTTAGATAGATCCAGTGGTTTAAACGGC
ACGTTTTTGCTTATAAAAAAAGTGCAAAAAAGATGTGGTTTACAAGTTAAAGCTACAGAA
TCCCTTTTTGCTGTAATTGCACCAGTTTTTAAAGCCTCTGGACAGAGCAGTATTTTCGTTTA
AACTTTGTTTTTCTTAAAAGCTTACAGTGTGTTGGCTAATTCTCCTCCCCTTTTTACAAG
ACGGGGGCGCGAGGGTGGACACTGGTGGCAGGTTAAGGGATACTGTCACTTTAAGAAGCC
25 TGCAGATTGAAGTGTAACATGGAGAAATTAGGGGCTGATTTTTTAACTGTGTGAGATA
TTAACCCAGCCGCCCTGTTATAAAATCAGGAAATCCAAACAGCGATTTACACCGATTAAACA
CCCCCTTTATATATTTTTTACAAAAATACACTGAGAAAAATAATCAAACGTTTTTCATCTCT
CTTGCTTTTTTTTTGTTTTTAAAAAGTGTCAAAAGTCTACATTTAAATATAAAAAATTTAAA
AGTTAAAACCTCTAGCCCTTCAGTGAAGGAGACGTAATGCGTGGGTAACAACAACACTAC
30 CAAAAAAGAAAGAAAAAAGAAAAAAGGAAAAGGAAGGAATAAAGAAATAAAGGAA
GTAAAAAGAAAGGAAAGAAAAAAGGGACAAAAGAAAAATATGTTTGGCCAGTATAAA
TACGTCCACATATAAAATGGCATCTGATTACATTTACAAGGAAAAAGAAATACGAGGAT
GGAGCATCGGTGAAGGAAAAAACACGTCTTCTCATTTACACCTATAAGGAATAAACACA
CACACTGAGAAAAAATTTGGCCCTGAATTGTTTTTTTTAAAGTCCAGCACAGATTTGAGTT
35 GCGTTTGAATCCTTTAAAGAGTTAAGAATGAAAAAAGCTGGTGATAATTTTCGTTGTAGG
AATCAAACATAGCGCCATCTATCTGCTTTTTTATATTATCCTACACTATTTTAAAAACTGC
TCAACAGTCTTATACAGAAATCTTAAAAGATAGACAGGATAACATGCTATATTAACCCC
ACCATTGAAATAATCCAACACCATCACGATTCCGATTAAGAGAAGNAAAAA

40 SEQ ID NO: 9

GAGAAATTTTAATTTGGTGTACAAAACCTTTAGTAGTTAGTGAATGAGTAAGATGATTTGTGTATTGTGCAG
TTTGGCTTTAAGAGCAAAGGAGAACTTAGAGACTATGGGTTCTAGATGCTGTGTCTGCCTTTCACAGTTTCT
CTGAACCTTTGTGATTGAACTGGATTTCATATTGGAGTTATATATCAGACAAAACCAAAGTCATTGGGCACAA
45 GCAGAATAAGTCCAAATTTAAAGTATAATCATGAGTTTGTAAATAATGCTTTATATACTTCTTTAAATAT
ACTTATTTAATATTAAGGATTTTTTTTTCTTGAATCCATATAATTTGAAAGGTCTAAGGGGAAAAACTTGT
TTTGAGCCACTTTATGATGTGAATTCACCTTCATAGGGAAATAACATGCTTGGTTTTTAATGTACTTTAACCC
TTTTAAAGTACACTGGACCAAATTTTAGAGAGCATATACATAGTGTTATAAAAT

50 SEQ ID NO: 10

TTTTTTTTTTTTAAGCTTTTGAGTGTTTTACACAAGGTATGTTACAGAAGACAAGTGCCTAAGTCAGAACTAT
GAGAAATATTGCCATCTGACAGGTTAGATTAAAGAAGGAGAGCACAGCATCTGGGAAGACTGAAAAGGGAAC
CATAGAAATAGGGGAAAAATGCTGAAACATGGGAGATGTCATTGAAGTCAAGAGAAAGAAAGCATCAGGAAA
55 ATTCAGGTGTCATGGAGAAAGCAAGTAAGATGGGGACAGTGAAGTTTGCTAGACATGATACAAGTAGATCGT
TAGGTGACTTACTGATACAACGTTATAAGTAGTGACCCCTTAACTGTAGGAGTTAATTCATACGTTGTAAAG
TTCTCCAGTATTGGTATTCATTGATCTTAATTCCTCTCTTTTTANAGAGNAAAGTTGCAGACGTGAATATGA
GGAGG

WO 00/43509

PCT/EP00/00346

SEQ ID NO: 11

CTAATACGACTCACTATAGGGCAGCGTGGTCGCGGCCGAGGTACATTTTTTATTACGTAACAGTAAAAACAT
 5 TAGAATAAGAAAAGACCTCCTTAATTCTGTTTCCCTTTCTTAAGGGGTAAGTATCATCTGCTCATTAACCTGA
 CAAATGGTGTGGTCAAGTAGCAAGGCTGCAGCCAAAAGACCTACACTGGATATCTGACTCCTTATCTTTCAT
 GTTGTGTGACCTTGAGCAAGTTACTTTAACCTTTTAAAGCCTCTGTTTTTTTCTTCTATAACAGAGACAGTA
 ATATTTATCACTGTTTTTTTTTGTAGAGATTAAATGGCGTATCTTATGTAAATCACCTAAAATATTTTTCTA
 10 GTAGTAGGGCATCAAAAATAAGCCCTGCTATTACTACTTTTTGAAGAAGATAAACCAAGTCCCTGGAACATTA
 ACCACCAAAAAGTACCACTGGTTTTTCACACCCCCC

SEQ ID NO: 12

CCTATGCATCAAGCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCAGTGTGCTGGAATTCGCCCTTA
 15 CCGACGTCGACTATCCAAGATGTACTATTATTGGCATATAGTTTCAAAAATTCACAGAAGGGAGCCAGGTGC
 TCTCATTGCCTAACAAAATAATGGAAATATGTATTCATTCTAACATCTTGACATACAGTATAAAGGGCACTC
 AGCAAGTGCTTTTTAGTTAGACTGATTTTAAATGAGTAGATTTCAGGAATACCAGAAAACGCACACACACACA
 GATAATTGGTTGAAAATACGAGGTTAGAACTTAACCATTTGAAGAGGAGTAAGGCTACTTAAAGCGTTAAAA
 ACTAATTGGGTAAGGTATGGTTGACCCAGCTACTTCATTGCGCTTACGATGTATATTCATTAAACTAAGTCA
 20 CTCACCTTCTCTGTGATGTTGATGGTTTGGTATAGTAAACATGAGATATGATTAAAGGTGATTTCAGGGATAG
 ATCAAGTGCTCGCCTAAGTAAATCTGGGNTTTCATTTTTTTCTAGACATCTTGGA

Claims

1. An isolated polynucleotide which comprises a nucleotide sequence which has at least 70% identity to that of SEQ ID NOs:1 - 6 over the entire length of SEQ ID NOs:1 - 6; or
5 a nucleotide sequence complementary to said isolated polynucleotide; and fragments of the polynucleotides.
2. The isolated polynucleotide as defined in claim 1 in which the identity is at least 95% to SEQ ID Nos: 1 - 6.
- 10 3. An isolated polynucleotide selected from:
 - (a) the polynucleotide of any one of SEQ ID NOs:1- 6; and
 - (b) a polynucleotide obtainable by screening an appropriate library under stringent hybridization conditions with a labeled probe having the sequence of any one of SEQ ID
15 NOs:1 - 6 or a fragment thereof; and
 - (c) a fragment of (a) or (b);or a nucleotide sequence complementary to said isolated polynucleotide
- 20 4. A process for diagnosing a disease or a susceptibility to a disease in a subject related to presence or expression or activity of a polynucleotide of any one of claims 1 to 3 in a subject comprising analyzing for the presence or amount of said polynucleotide in a sample derived from said subject.
- 25 5. A process for diagnosing the presence of colon cancer or a susceptibility to colon cancer in a subject related to presence or expression or activity of a polynucleotide of any one of claims 1 to 3 in a subject comprising analyzing for the presence or amount of said polynucleotide in a sample derived from said subject.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C12N 15/11, 15/12, C12Q 1/68		A2	(11) International Publication Number: WO 00/43509
			(43) International Publication Date: 27 July 2000 (27.07.00)
(21) International Application Number: PCT/EP00/00346		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: 17 January 2000 (17.01.00)			
(30) Priority Data: 9901078.7 19 January 1999 (19.01.99) GB 9902090.1 29 January 1999 (29.01.99) GB 9902169.3 1 February 1999 (01.02.99) GB 9902168.5 1 February 1999 (01.02.99) GB 9902163.6 1 February 1999 (01.02.99) GB 9907901.4 7 April 1999 (07.04.99) GB			
(71) Applicant (for all designated States except US): SMITHKLINE BEECHAM BIOLOGICALS S.A. [BE/BE]; Rue de l'Institut 89, B-1330 Rixensart (BE).		Published <i>Without international search report and to be republished upon receipt of that report.</i>	
(72) Inventor; and (75) Inventor/Applicant (for US only): VINALS-BASSOLS, Carlota [BE/BE]; SmithKline Beecham Biologicals S.A., Rue de l'Institut 89, B-1330 Rixensart (BE).			
(74) Agent: TYRRELL, Arthur, William, Russell; SmithKline Beecham, Two New Horizons Court, Brentford, Middlesex TW8 9EP (GB).			
(54) Title: NOVEL COMPOUNDS			
(57) Abstract <p>CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptides and polynucleotides and methods for producing such polypeptides by recombinant techniques are disclosed. Also disclosed are methods for utilizing CASB611, CASB500, CASB501, CASB502, CASB505, or CASB507 polypeptides and polynucleotides in diagnostics, and vaccines for prophylactic and therapeutic treatment of cancers, particularly colon cancers, autoimmune diseases, and related conditions.</p>			

Docket No.: BC45215

PCT/EP00/00346

DECLARATION AND POWER OF ATTORNEY

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

POLYNUCLEOTIDES AND THEIR APPLICATIONS IN DIAGNOSTICS

the specification of which (check one)

☐ is attached hereto.

☒ was filed on 17 January 2000 as Serial No. PCT/EP00/00346
and was amended on (if applicable).

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the patentability as defined in Title 37, Code of Federal Regulations, Section 1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, Section 119(a)-(d) or Section 365(b) of any foreign application(s) for patent or inventor's certificate, or Section 365(a) of any PCT International application which designated at least one country other than the United States, listed below and have also identified below any foreign application for patent or Inventor's certificate, or PCT International application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application(s)

Number	Country	Filing Date	Priority Claimed
9901078.7	Great Britain	19 January 1999	Yes
9902090.1	Great Britain	29 January 1999	Yes
9902169.3	Great Britain	1 February 1999	Yes
9902168.5	Great Britain	1 February 1999	Yes
9902163.6	Great Britain	1 February 1999	Yes
9907901.4	Great Britain	7 April 1999	Yes

I hereby claim the benefit under Title 35, United States Code, Section 119(e) of any United States provisional application(s) listed below.



Application Number Filing Date

I hereby claim the benefit under Title 35, United States Code, Section 120 of any United States application(s) or Section 365(c) of any PCT International application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of Title 35, United States Code, Section 112, I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations, Section 1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application.

Serial No. Filing Date Status

I hereby appoint the practitioners associated with the Customer Number provided below to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith, and direct that all correspondence be addressed to that Customer Number:

Customer Number 25,308

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full Name of Inventor: ¹⁻⁰⁰ Carlota VINALS-BASSOLS

Inventor's Signature: Carlota

Date: Aug 1st 2001

Residence: Rixensart, Belgium

BEL



Citizenship: Belgian

**Post Office Address: SmithKline Beecham Corporation
Corporate Intellectual Property - UW2220
P.O. Box 1539
King of Prussia, Pennsylvania 19406-0939**